

10/718480  
STN Search Summary

=> d his

(FILE 'HOME' ENTERED AT 15:22:48 ON 06 OCT 2004)

FILE 'REGISTRY' ENTERED AT 15:22:57 ON 06 OCT 2004

L1 0 S LYSE/CN  
L2 92 S LYSE  
L3 7 S L2 AND LYSINE

FILE 'CAPLUS' ENTERED AT 15:24:09 ON 06 OCT 2004

L4 10 S L3  
L5 1 S L4 AND (METHYLBACILL? OR METHYLOPHIL? OR FLAGELLAT? OR GLYCOG  
L6 1 S L4 AND (METHANOL? OR METHYL?)

L3 ANSWER 1 OF 7 REGISTRY COPYRIGHT 2004 ACS on STN  
RN 615223-10-4 REGISTRY  
CN Lysine exporter protein (Corynebacterium diphtheriae strain NCTC13129  
gene lyse) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank CAE49614  
CN GenBank CAE49614 (Translated from: GenBank BX248357)

L3 ANSWER 2 OF 7 REGISTRY COPYRIGHT 2004 ACS on STN  
RN 571234-77-0 REGISTRY  
CN Lysine exporter protein (Corynebacterium efficiens strain YS-314 gene  
lyse) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank BAC18167  
CN GenBank BAC18167 (Translated from: GenBank AP005218)

L3 ANSWER 3 OF 7 REGISTRY COPYRIGHT 2004 ACS on STN  
RN 478431-27-5 REGISTRY  
CN 1-125-Protein (Corynebacterium lactofermentum gene lyse) (9CI)  
(CA INDEX NAME)

OTHER NAMES:

CN 9: PN: EP1266966 SEQID: 10 claimed protein  
CN Protein (Corynebacterium lactofermentum gene lyse24  
lysine-exporting)  
CN Transport protein (Corynebacterium lactofermentum gene lyse24  
lysine-exporting)

L3 ANSWER 4 OF 7 REGISTRY COPYRIGHT 2004 ACS on STN  
RN 338802-47-4 REGISTRY  
CN Lysine-transporting protein (Corynebacterium thermoaminogenes strain  
AJ12310 gene lyse) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 16: PN: JP2001120270 SEQID: 14 claimed protein

L3 ANSWER 5 OF 7 REGISTRY COPYRIGHT 2004 ACS on STN  
 RN 314317-10-7 REGISTRY  
 CN Protein MP (metabolic pathway) (Corynebacterium glutamicum strain ATCC\_13032 clone RXC01796) (9CI) (CA INDEX NAME)  
 OTHER NAMES:  
 CN 10: PN: WO0100843 SEQID: 10 claimed protein  
 CN 16: PN: WO0166573 SEQID: 16 claimed protein  
 CN Lysine-transporting protein (Corynebacterium glutamicum strain ATCC10032 clone RXC01796 gene lysE)

L3 ANSWER 6 OF 7 REGISTRY COPYRIGHT 2004 ACS on STN  
 RN 314317-08-3 REGISTRY  
 CN Protein MP (metabolic pathway) (Corynebacterium glutamicum strain ATCC\_13032 clone RXC02390) (9CI) (CA INDEX NAME)  
 OTHER NAMES:  
 CN 14: PN: WO0166573 SEQID: 14 claimed protein  
 CN 8: PN: WO0100843 SEQID: 8 claimed protein  
 CN Lysine-transporting protein (Corynebacterium glutamicum strain ATCC13032 clone RXC02390 gene lysE)

L3 ANSWER 7 OF 7 REGISTRY COPYRIGHT 2004 ACS on STN  
 RN 184922-77-8 REGISTRY  
 CN Lysine-transporter (Corynebacterium glutamicum strain R127 gene lysE) (9CI) (CA INDEX NAME)  
 OTHER NAMES:  
 CN 3454: PN: EP1108790 SEQID: 6955 claimed protein  
 CN Lysine exporter protein (Corynebacterium glutamicum strain R127 lysE gene)  
 CN Lysine-transporting protein (Corynebacterium glutamicum strain R127 lysE gene)  
 CN Protein (Corynebacterium glutamicum strain ATCC13032 clone EP1108790-SEQID-6955)  
 CN Protein (Corynebacterium glutamicum strain R127 gene lysE)

L4 ANSWER 1 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 2003:904048 CAPLUS  
 TI The complete genome sequence and analysis of Corynebacterium diphtheriae NCTC13129  
 AU Cerdano-Tarraga, A. M.; Efstratiou, A.; Dover, L. G.; Holden, M. T. G.; Pallen, M.; Bentley, S. D.; Besra, G. S.; Churcher, C.; James, K. D.; De Zoysa, A.; Chillingworth, T.; Cronin, A.; Dowd, L.; Feltwell, T.; Hamlin, N.; Holroyd, S.; Jagels, K.; Moule, S.; Quail, M. A.; Rabinowitsch, E.; Rutherford, K. M.; Thomson, N. R.; Unwin, L.; Whitehead, S.; Barrell, B. G.; Parkhill, J.  
 SO Nucleic Acids Research (2003), 31(22), 6516-6523

L4 ANSWER 2 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 2003:566882 CAPLUS  
 TI Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of Corynebacterium efficiens  
 AU Nishio, Yousuke; Nakamura, Yoji; Kawarabayashi, Yutaka; Usuda, Yoshihiro; Kimura, Eiichiro; Sugimoto, Shinichi; Matsui, Kazuhiko; Yamagishi, Akihiko; Kikuchi, Hisashi; Ikeo, Kazuho; Gojobori, Takashi  
 SO Genome Research (2003), 13(7), 1572-1579

★  
14/04/03

L4 ANSWER 3 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 2002:963758 CAPLUS  
 TI L-lysine or L-arginine fermentation using a methanol assimilating  
 bacterium which secretes amino acids  
 IN Gunji, Yoshiya; Yasueda, Hisashi

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 1266966	A2	20021218	EP 2002-12539	20020605
EP 1266966	A3	20040114		
JP 2003061687	A2	20030304	JP 2002-151981	20020527
<b>US 2003124687</b>	<b>A1</b>	<b>20030703</b>	US 2002-166142	20020611
PRAI JP 2001-177075	A	20010612		

L4 ANSWER 4 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 2001:676795 CAPLUS  
 TI Corynebacterium glutamicum genes encoding metabolic pathway proteins  
 IN Pompejus, Markus; Kroeger, Burkhard; Schroeder, Hartwig; Zelder, Oskar;  
 Haberhauer, Gregor; Kim, Jun-Won; Lee, Heung-Shick; Hwang, Byung-Joon

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001066573	A2	20010913	WO 2000-IB2035	20001222
WO 2001066573	A3	20020510		
EP 1261718	A2	20021204	EP 2000-987602	20001222
BR 2000017148	A	20030311	BR 2000-17148	20001222
JP 2003525623	T2	20030902	JP 2001-565737	20001222
ZA 2002008060	A	20031110	ZA 2002-8060	20021008
PRAI US 2000-187970P	P	20000309		
US 2000-606740	A	20000623		
WO 2000-IB2035	W	20001222		

L4 ANSWER 5 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 2001:574896 CAPLUS  
 Correction of: 2001:450980  
 DN 135:283981  
 Correction of: 135:29919  
 TI Complete genome sequence of Corynebacterium glutamicum ATCC 13032 and its  
 genes and encoded proteins  
 IN Nakagawa, Satoshi; Mizoguchi, Hiroshi; Ando, Seiko; Hayashi, Mikiro;  
 Ochiai, Keiko; Yokoi, Haruhiko; Tateishi, Naoko; Senoh, Akihiro; Ikeda,  
 Masato; Ozaki, Akio

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 1108790	A2	20010620	EP 2000-127688	20001218
JP 2002191370	A2	20020709	JP 2000-405096	20001215
US 2002197605	A1	20021226	US 2000-738626	20001218
PRAI JP 1999-377484	A	19991216		
JP 2000-159162	A	20000407		
JP 2000-280988	A	20000803		

L4 ANSWER 6 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 2001:324284 CAPLUS  
 TI Corynebacterium thermoaminogenes thermostable L-lysine biosynthesis genes  
 IN Itaya, Hiroshi; Kimura, Eiichiro; Kawahara, Yoshio; Sugimoto, Shinichi  
 LA Japanese

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	JP 2001120270	A2	20010508	JP 1999-311148	19991101
PRAI	JP 1999-311148		19991101		

L4 ANSWER 7 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 2001:12637 CAPLUS  
 TI Corynebacterium glutamicum genes encoding proteins involved in metabolic pathways  
 IN Pompejus, Markus; Kroger, Burkhard; Schroder, Hartwig; Zelder, Oskar; Haberhauer, Gregor

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2001000843	A2	20010104	WO 2000-IB923	20000623
	BR 2000011806	A	20020514	BR 2000-11806	20000623
	TR 200103707	T2	20020923	TR 2001-200103707	20000623
	EP 1257649	A2	20021120	EP 2000-938999	20000623
	US 2004030116	A1	20040212	US 2003-627476	20030725
	US 2004180408	A1	20040916	US 2004-781014	20040217
PRAI	US 1999-141031P	P	19990625		
	DE 1999-19930476	A	19990701		
	US 1999-142101P	P	19990702		
	DE 1999-19931415	A	19990708		
	DE 1999-19931418	A	19990708		
	DE 1999-19931419	A	19990708		
	DE 1999-19931420	A	19990708		
	DE 1999-19931424	A	19990708		
	DE 1999-19931428	A	19990708		
	DE 1999-19931434	A	19990708		
	DE 1999-19931435	A	19990708		
	DE 1999-19931443	A	19990708		
	DE 1999-19931453	A	19990708		
	DE 1999-19931457	A	19990708		
	DE 1999-19931465	A	19990708		
	DE 1999-19931478	A	19990708		
	DE 1999-19931510	A	19990708		
	DE 1999-19931541	A	19990708		
	DE 1999-19931573	A	19990708		
	DE 1999-19931592	A	19990708		
	DE 1999-19931632	A	19990708		
	DE 1999-19931634	A	19990708		
	DE 1999-19931636	A	19990708		
	DE 1999-19932125	A	19990709		
	DE 1999-19932126	A	19990709		
	DE 1999-19932130	A	19990709		
	DE 1999-19932186	A	19990709		
	DE 1999-19932206	A	19990709		
	DE 1999-19932227	A	19990709		
	DE 1999-19932228	A	19990709		
	DE 1999-19932229	A	19990709		
	DE 1999-19932230	A	19990709		
	DE 1999-19932922	A	19990714		

DE 1999-19932926	A	19990714
DE 1999-19932928	A	19990714
DE 1999-19933004	A	19990714
DE 1999-19933005	A	19990714
DE 1999-19933006	A	19990714
US 1999-148613P	P	19990812
DE 1999-19940764	A	19990827
DE 1999-19931412	A	19990708
DE 1999-19931413	A	19990708
DE 1999-19931431	A	19990708
DE 1999-19931433	A	19990708
DE 1999-19931454	A	19990708
DE 1999-19931562	A	19990708
DE 1999-19931563	A	19990708
DE 1999-19932122	A	19990709
DE 1999-19932124	A	19990709
DE 1999-19932128	A	19990709
DE 1999-19932180	A	19990709
DE 1999-19932182	A	19990709
DE 1999-19932190	A	19990709
DE 1999-19932191	A	19990709
DE 1999-19932209	A	19990709
DE 1999-19932212	A	19990709
US 1999-143208P	P	19990709
DE 1999-19932924	A	19990714
DE 1999-19932927	A	19990714
DE 1999-19932973	A	19990714
DE 1999-19940765	A	19990827
DE 1999-19940766	A	19990827
DE 1999-19940830	A	19990827
DE 1999-19940831	A	19990827
DE 1999-19940832	A	19990827
DE 1999-19940833	A	19990827
DE 1999-19941378	A	19990831
DE 1999-19941379	A	19990831
DE 1999-19941395	A	19990831
US 1999-151572P	P	19990831
DE 1999-19942076	A	19990903
DE 1999-19942077	A	19990903
DE 1999-19942078	A	19990903
DE 1999-19942079	A	19990903
DE 1999-19942086	A	19990903
DE 1999-19942087	A	19990903
DE 1999-19942088	A	19990903
DE 1999-19942095	A	19990903
DE 1999-19942123	A	19990903
DE 1999-19942125	A	19990903
US 2000-602740	A1	20000623
US 2000-602787	A1	20000623
WO 2000-IB923	W	20000623

L4 ANSWER 8 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 1997:500174 CAPLUS  
 TI Microorganism amino acid exporter or gene for use in amino acid production by fermentation  
 IN Vrljic, Marina; Eggeling, Lothar; Sahm, Hermann  
 LA German

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	DE 19548222	A1	19970626	DE 1995-19548222	19951222
	CA 2241111	AA	19970703	CA 1996-2241111	19961218
	WO 9723597	A2	19970703	WO 1996-DE2485	19961218
	WO 9723597	A3	19971023		
	AU 9719218	A1	19970717	AU 1997-19218	19961218
	AU 724536	B2	20000921		
	EP 868527	A2	19981007	EP 1996-946220	19961218
	CN 1209169	A	19990224	CN 1996-180096	19961218
	BR 9612666	A	19991005	BR 1996-12666	19961218
	JP 2000507086	T2	20000613	JP 1997-523222	19961218
	RU 2225883	C2	20040320	RU 1998-113861	19961218
	ZA 9610768	A	19980731	ZA 1996-10768	19961220
PRAI	DE 1995-19548222	A	19951222		
	WO 1996-DE2485	W	19961218		

L4 ANSWER 9 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 1997:475788 CAPLUS  
 TI L-Lysine export from Corynebacterium glutamicum. Physiological and molecular-biological characterization of the carrier-mediated export of a primary metabolite  
 AU Vrljic, Maria-Marina  
 SO Berichte des Forschungszentrums Juelich (1997), Juel-3349, 1-115 pp.  
 LA German

L4 ANSWER 10 OF 10 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 1996:757064 CAPLUS  
 TI A new type of transporter with a new type of cellular function: L-lysine export from Corynebacterium glutamicum  
 AU Vrljic, Marina; Sham, Hermann; Eggeling, Lothar  
 SO Molecular Microbiology (1996), 22(5), 815-826

10/716480  
STN Search Summary 2

=> d his

FILE 'CAPLUS' ENTERED AT 17:35:10 ON 06 OCT 2004

L1 2182702 (METHYLBACILL? OR METHYLOPHIL? OR FLAGELLAT? OR GLYCOGENE? OR  
METHANOL? OR METHYL?)  
L2 96 L1 AND LYSE  
L3 6 L2 AND LYSINE  
L4 1 L1 AND LYSINE(2W)EXPORT  
L5 58 L1 AND LYSINE(2W)TRANSPORT  
L6 5 L1 AND LYSINE(2W)RESISTANCE  
L7 1 L5 AND LYSINE(2W)ANALOG?

L4 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2004 ACS on STN  
AN 2002:301460 CAPLUS  
TI Lipid synthesis in Corynebacterium glutamicum: Genetical and biochemical  
investigations of acyl-CoA carboxylases  
AU Tilg, Yvonne  
CS Institut fur Biotechnologie, Germany  
SO Berichte des Forschungszentrums Juelich (2002), Juel-3946, i-ix, 1-133  
CODEN: FJBEE5; ISSN: 0366-0885  
LA German

L6 ANSWER 1 OF 5 CAPLUS COPYRIGHT 2004 ACS on STN  
AN 1995:722218 CAPLUS  
TI Selection and characterization of aspartokinase feedback-insensitive  
mutants of Azotobacter vinelandii  
AU Ekechukwu, Chioma R.; Burns, Thomas A.; Melton, Thoyd  
CS Dep. Microbiology, North Carolina State Univ., Raleigh, NC, 27695, USA  
SO Applied and Environmental Microbiology (1995), 61(8), 3189-91  
CODEN: AEMIDF; ISSN: 0099-2240

L6 ANSWER 2 OF 5 CAPLUS COPYRIGHT 2004 ACS on STN  
AN 1993:144831 CAPLUS  
TI Endothelium-dependent relaxation of hypertensive resistance arteries is  
not impaired under all conditions  
AU Li, Junyi; Bukoski, Richard D.  
CS Med. Branch, Univ. Texas, Galveston, TX, 77550, USA  
SO Circulation Research (1993), 72(2), 290-6  
CODEN: CIRUAL; ISSN: 0009-7330

L6 ANSWER 3 OF 5 CAPLUS COPYRIGHT 2004 ACS on STN  
AN 1991:629291 CAPLUS  
TI Isolation of variants of Chinese hamster ovary cells with abnormally low  
levels of GSH: decreased ability to cleave endocytosed disulfide bonds  
AU Mandel, Richard; Ryser, Hugues J. P.; Niaki, Bijan; Ghani, Farooq; Shen,  
Wei Chiang  
CS Sch. Med., Boston Univ., Boston, MA, 02118, USA  
SO Journal of Cellular Physiology (1991), 149(1), 60-5  
CODEN: JCLLAX; ISSN: 0021-9541

L6 ANSWER 4 OF 5 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 1991:120274 CAPLUS  
 TI Fermentative manufacture of L-lysine with Brevibacterium or  
 Corynebacterium species  
 IN Yoshihara, Yasuhiko; Kawahara, Yoshio; Ishii, Toshimasa  
 PA Ajinomoto Co., Inc., Japan  
 SO Jpn. Kokai Tokkyo Koho, 4 pp.  
 CODEN: JKXXAF  
 DT Patent  
 LA Japanese  
 FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
	-----	----	-----	-----	-----
PI	JP 02234686	A2	19900917	JP 1989-54192	19890307
	JP 2817172	B2	19981027		
	CN 1045419	A	19900919	CN 1990-101212	19900307
	CN 1030616	B	19960103		
PRAI	JP 1989-54192	A	19890307		

L6 ANSWER 5 OF 5 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 1988:526016 CAPLUS  
 TI Selection of regenerable maize callus cultures resistant to 5-  
 methyl-DL-tryptophan, S-2-aminoethyl-L-cysteine and high levels of  
 L-lysine plus L-threonine  
 AU Miao, Shuhua; Duncan, David R.; Widholm, Jack  
 CS Dep. Agron., Univ. Illinois, Urbana, IL, 61801, USA  
 SO Plant Cell, Tissue and Organ Culture (1988), 14(1), 3-14  
 CODEN: PTCEDJ; ISSN: 0167-6857

L7 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2004 ACS on STN  
 AN 1977:167623 CAPLUS  
 TI Characterization of a lysine-specific active transport  
 system in Rickettsia prowazeki  
 AU Smith, Deborah K.; Winkler, Herbert H.  
 CS Sch. Med., Univ. Virginia, Charlottesville, VA, USA  
 SO Journal of Bacteriology (1977), 129(3), 1349-55  
 CODEN: JOBAAY; ISSN: 0021-9193

=> d 17 abs

L7 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2004 ACS on STN  
 AB R. prowazeki possesses an active transport system for lysine with a Kt of  
 influx of 1 .mu.M. This intracellular lysine pool can be exchanged with  
 external unlabeled substrates for a least 10 min. The lysine  
 analogs L-aminoethyl cysteine, N-methyl lysine,  
 hydroxylysine, and D-lysine competitively inhibited uptake of L-lysine,  
 but cadaverine, diaminopimelate, arginine, ornithine, and  
 .epsilon.-aminocaproate did not. Accumulation of lysine was inhibited by  
 the energy poisons KCN, triphenylmethyl phosphonium Br, and  
 2,4-dinitrophenol. The effect of KCN, but not 2,4-dinitrophenol or  
 triphenylmethyl phosphonium Br, was overcome by ATP. Both  
 energy-dependent influx and efflux were inhibited by the sulfhydryl  
 reagents N-ethyl maleimide and p-chloromercuriphenyl sulfonic acid.

# ~~STAT~~ SEQ SEARCH SUMMARY

10/716480

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2004, 11:16:40 ; Search time 2934 Seconds  
(without alignments)  
10503.373 Million cell updates/sec

Title: US-10-716-480A-1  
Perfect score: 711  
Sequence: 1 atggtgatcatggaaatctt.....aactgatgttgatgggtag 711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%					Description
	No.	Score	Match	Length	ID	
	1	711	100.0	711	6 AX643028	AX643028 Sequence
	2	711	100.0	822	6 AX063771	AX063771 Sequence
	3	711	100.0	822	6 AX244059	AX244059 Sequence
	4	711	100.0	2374	1 CGLYSEG	X96471 C.glutamicu
	5	711	100.0	2374	6 A93933	A93933 Sequence 2
c	6	711	100.0	333150	1 AP005277	AP005277 Corynebac
c	7	711	100.0	349980	6 AX127147	AX127147 Sequence
	8	708	99.6	708	6 AX123539	AX123539 Sequence
	9	708	99.6	708	6 BD165656	BD165656 Novel pol
	10	696.8	98.0	712	6 AX643030	AX643030 Sequence
	11	317.6	44.7	1568	6 E54483	E54483 Heat-resist
	12	317.6	44.7	1771	1 AB083133	AB083133 Corynebac
c	13	317.6	44.7	308650	1 AP005218	AP005218 Corynebac
c	14	136.6	19.2	349535	1 BX248357	BX248357 Corynebac
	15	84.2	11.8	190050	1 AL646059	AL646059 Ralstonia
c	16	65.8	9.3	300861	1 AE016777	AE016777 Pseudomon
	17	65.6	9.2	303550	1 SCO939131	AL939131 Streptomy
	18	63.4	8.9	10542	1 AE004852	AE004852 Pseudomon
c	19	63.2	8.9	10024	1 AE015725	AE015725 Shewanell
c	20	62.8	8.8	24336	1 AE008841	AE008841 Salmonell
c	21	62.8	8.8	298300	1 AP005025	AP005025 Streptomy
	22	62.2	8.7	311600	1 AE016871	AE016871 Pseudomon
c	23	62	8.7	10313	1 AE015304	AE015304 Shigella
c	24	62	8.7	10701	1 AE005522	AE005522 Escherich
c	25	62	8.7	266658	1 AP002563	AP002563 Escherich
c	26	62	8.7	292906	1 AE016988	AE016988 Shigella
	27	61.6	8.7	4033	1 ASU65741	U65741 Aeromonas s
c	28	61.2	8.6	230050	1 AL627277	AL627277 Salmonell
c	29	61.2	8.6	301574	1 AE016844	AE016844 Salmonell
	30	60.4	8.5	636	6 E49392	E49392 Process for
	31	60.4	8.5	636	6 AX030085	AX030085 Sequence
	32	60.4	8.5	8029	1 ECFDAPGK	X14436 Escherichia
c	33	60.4	8.5	10362	1 AE000375	AE000375 Escherich

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2004, 09:28:59 ; Search time 355 Seconds  
(without alignments)  
8508.363 Million cell updates/sec

Title: US-10-716-480A-1  
Perfect score: 711  
Sequence: 1 atggtgatcatggaaatctt.....aactgatgttgatgggtag 711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	711	100.0	711	8	ACC80941	Acc80941 LysE prot
2	711	100.0	822	4	AAF71779	Aaf71779 Corynebac
3	711	100.0	822	4	AAS96098	Aas96098 C. glutam
4	711	100.0	2374	2	AAT96816	Aat96816 DNA encod
5	711	100.0	2374	9	ADB66196	Adb66196 DNA fragm
c 6	711	100.0	349980	5	AAH68528	Aah68528 C glutami
7	708	99.6	708	5	AAH68420	Aah68420 C glutami
8	696.8	98.0	712	8	ACC80942	Acc80942 LysE24 pr
9	317.6	44.7	1568	4	AAH45375	Aah45375 C. thermo
10	60.4	8.5	636	3	AAA52691	Aaa52691 Escherich

OM nucleic - nucleic search, using sw model

Run on: October 7, 2004, 12:21:55 ; Search time 83 Seconds  
(without alignments)  
4753.856 Million cell updates/sec

Title: US-10-716-480A-1  
Perfect score: 711  
Sequence: 1 atggtgatcatggaaatctt.....aactgatgttgatgggtag 711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result			%				ID	Description
	No.	Score	Match	Length	DB	Query		
c	1	61.8	8.7	699	4	US-09-252-991A-7908		Sequence 7908, Ap
	2	61.8	8.7	834	4	US-09-252-991A-7643		Sequence 7643, Ap
	3	61.8	8.7	894	4	US-09-252-991A-7841		Sequence 7841, Ap
c	4	57	8.0	597	4	US-09-894-844-11		Sequence 11, Appl
	5	57	8.0	15239	1	US-08-390-878-17		Sequence 17, Appl
	6	57	8.0	4403765	3	US-09-103-840A-2		Sequence 2, Appli
c	7	57	8.0	4411529	3	US-09-103-840A-1		Sequence 1, Appli
	8	56.6	8.0	642	4	US-09-489-039A-4674		Sequence 4674, Ap
	9	53.8	7.6	1176	4	US-09-252-991A-7571		Sequence 7571, Ap
c	10	43	6.0	645	4	US-09-543-681A-800		Sequence 800, App
	11	39.4	5.5	505	4	US-09-621-976-15639		Sequence 15639, A
	12	37.6	5.3	4403765	3	US-09-103-840A-2		Sequence 2, Appli

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2004, 12:38:35 ; Search time 428 Seconds  
(without alignments)  
8421.640 Million cell updates/sec

Title: US-10-716-480A-1  
Perfect score: 711  
Sequence: 1 atggtgatcatggaaatctt.....aactgatgttgatgggtag 711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result		%	Query				
No.	Score	Match	Length	DB	ID	Description	
1	711	100.0	711	15	US-10-166-142-7	Sequence 7, Appli	
2	711	100.0	711	17	US-10-716-473-7	Sequence 7, Appli	
3	711	100.0	822	10	US-09-746-660A-51	Sequence 51, Appl	
4	711	100.0	2374	15	US-10-196-232-24	Sequence 24, Appl	
c 5	711	100.0	3309400	9	US-09-738-626-1	Sequence 1, Appli	
6	708	99.6	708	9	US-09-738-626-3455	Sequence 3455, Ap	
7	696.8	98.0	712	15	US-10-166-142-9	Sequence 9, Appli	
8	696.8	98.0	712	17	US-10-716-473-9	Sequence 9, Appli	
9	62.8	8.8	609	15	US-10-156-761-1109	Sequence 1109, Ap	
c 10	62.8	8.8	9025608	15	US-10-156-761-1	Sequence 1, Appli	
11	57	8.0	597	9	US-09-894-844-11	Sequence 11, Appl	
12	57	8.0	597	13	US-10-647-089-11	Sequence 11, Appl	
13	57	8.0	597	16	US-10-388-902-11	Sequence 11, Appl	
14	56.2	7.9	86114	15	US-10-080-170-648	Sequence 648, App	
15	56.2	7.9	86114	17	US-10-080-170-648	Sequence 648, App	
c 16	42.2	5.9	1780	17	US-10-437-963-18491	Sequence 18491, A	
17	42.2	5.9	3000	15	US-10-156-761-5949	Sequence 5949, Ap	
18	42.2	5.9	9025608	15	US-10-156-761-1	Sequence 1, Appli	
19	42	5.9	2301	17	US-10-437-963-6293	Sequence 6293, Ap	
20	41.2	5.8	2016	17	US-10-437-963-86805	Sequence 86805, A	
21	40.4	5.7	1392	16	US-10-369-493-31368	Sequence 31368, A	
22	40.4	5.7	1443	16	US-10-369-493-28610	Sequence 28610, A	
23	40.4	5.7	2429	17	US-10-437-963-6292	Sequence 6292, Ap	
c 24	40	5.6	3483	17	US-10-437-963-44614	Sequence 44614, A	
c 25	39.4	5.5	594	16	US-10-369-493-33277	Sequence 33277, A	
c 26	39.2	5.5	752	17	US-10-437-963-23986	Sequence 23986, A	
c 27	39.2	5.5	1257	15	US-10-156-761-5999	Sequence 5999, Ap	
c 28	39	5.5	993	10	US-09-746-660A-47	Sequence 47, Appl	
29	39	5.5	2823	13	US-10-282-122A-26243	Sequence 26243, A	
30	39	5.5	2826	13	US-10-282-122A-28428	Sequence 28428, A	
31	38.6	5.4	1245	13	US-10-282-122A-14346	Sequence 14346, A	
32	38.4	5.4	1236	17	US-10-437-963-95996	Sequence 95996, A	
c 33	38.2	5.4	1365	17	US-10-437-963-14494	Sequence 14494, A	
c 34	38	5.3	1018	13	US-10-425-114-960	Sequence 960, App	
35	38	5.3	1458	15	US-10-156-761-1306	Sequence 1306, Ap	
c 36	38	5.3	2218	13	US-10-425-114-28221	Sequence 28221, A	
c 37	38	5.3	2457	13	US-10-425-114-30614	Sequence 30614, A	
38	37.8	5.3	270	17	US-10-437-963-34340	Sequence 34340, A	
c 39	37.8	5.3	536	17	US-10-338-110-119	Sequence 119, App	
40	37.8	5.3	2038	15	US-10-094-113-11	Sequence 11, Appl	
41	37.6	5.3	536	17	US-10-338-110-119	Sequence 119, App	
c 42	37.6	5.3	2079	15	US-10-156-761-1755	Sequence 1755, Ap	
c 43	37.6	5.3	3073	13	US-10-424-599-33204	Sequence 33204, A	
c 44	37.6	5.3	3090	13	US-10-425-114-11304	Sequence 11304, A	
45	37.4	5.3	624	15	US-10-156-761-2439	Sequence 2439, Ap	

## ALIGNMENTS

RESULT 1  
US-10-166-142-7



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2004, 12:24:20 ; Search time 2516 Seconds  
(without alignments)  
8438.796 Million cell updates/sec

Title: US-10-716-480A-1  
Perfect score: 711  
Sequence: 1 atggtgatcatggaaatctt.....aactgatgttgatgggtag 711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame plus p2n model

```
Run on:      October  7, 2004, 15:19:08 ; Search time 2849 Seconds
              (without alignments)
              3590.367 Million cell updates/sec
```

Title: US-10-716-480A-2  
Perfect score: 1191  
Sequence: 1 MVIMEIFITGLLLGASLLLS.....INVVVAVVMTALAIKLMLMG 236

```
Scoring table:  BLOSUM62
                  Xgapop 10.0 , Xgapext 0.5
                  Ygapop 10.0 , Ygapext 0.5
                  Fgapop  6.0 , Fgapext 7.0
                  Delop  6.0 , Delext 7.0
```

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
```

Q=/cgn2\_1/USPTO\_spool\_p/US10716480/runat\_06102004\_102309\_20226/app\_query.fasta\_1  
.391

```
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716480_CGN_1_1_5600@runat_06102004_102309_20226 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

```
Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
```

```

11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rad:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1191	100.0	708	6	AX123539	Sequence
2	1191	100.0	708	6	BD165656	Novel pol
3	1191	100.0	711	6	AX643028	Sequence
4	1191	100.0	822	6	AX063771	Sequence
5	1191	100.0	822	6	AX244059	Sequence
6	1191	100.0	2374	1	CGLYSEG	C.glutamicu
7	1191	100.0	2374	6	A93933	Sequence 2
c 8	1191	100.0	333150	1	AP005277	Corynebac
c 9	1191	100.0	349980	6	AX127147	Sequence
10	1169	98.2	712	6	AX643030	Sequence
11	864.5	72.6	1568	6	E54483	Heat-resist
12	864.5	72.6	1771	1	AB083133	Corynebac
c 13	864.5	72.6	308650	1	AP005218	Corynebac
c 14	562	47.2	349535	1	BX248357	Corynebac
15	352	29.6	636	6	E49392	Process for
16	352	29.6	636	6	AX030085	Sequence

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 7, 2004, 14:30:02 ; Search time 340 Seconds  
(without alignments)  
2948.749 Million cell updates/sec

Title: US-10-716-480A-2  
Perfect score: 1191  
Sequence: 1 MVIMEIFITGLLLGLASLLLS.....INVVVAVVMTALAIKMLMG 236

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delect 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp

Q=/cgn2\_1/USPTO\_spool\_p/US10716480/runat\_06102004\_102309\_20217/app\_query.fasta\_1  
.391

-DB=N\_Geneseq\_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10716480\_@CGN\_1\_1\_708\_@runat\_06102004\_102309\_20217 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1191	100.0	708	5	AAH68420	Aah68420 C glutami
2	1191	100.0	711	8	ACC80941	Acc80941 LysE prot
3	1191	100.0	822	4	AAF71779	Aaf71779 Corynebac
4	1191	100.0	822	4	AAS96098	Aas96098 C. glutam
5	1191	100.0	2374	2	AAT96816	Aat96816 DNA encod
6	1191	100.0	2374	9	ADB66196	Adb66196 DNA fragm
c 7	1191	100.0	349980	5	AAH68528	Aah68528 C glutami
8	1169	98.2	712	8	ACC80942	Acc80942 LysE24 pr
9	864.5	72.6	1568	4	AAH45375	Aah45375 C. thermo
* 10	352	29.6	636	3	AAA52691	Aaa52691 Escherich
11	327.5	27.5	86114	6	ABX09143	Abx09143 Mycobacte
12	327.5	27.5	110000	4	AAI99682_05	Continuation (6 of
13	327.5	27.5	110000	4	AAI99683_05	Continuation (6 of
c 14	318	26.7	15239	2	AAT33536	Aat33536 BCG delet
15	318	26.7	110000	4	AAI99682_22	Continuation (23 o
16	318	26.7	110000	4	AAI99683_22	Continuation (23 o
17	298	25.0	624	7	ACF71727	Acf71727 Photorhab
18	298	25.0	636	8	ADA30324	Ada30324 DNA encod
19	298	25.0	110000	7	ACF67367_49	Continuation (50 o
20	298	25.0	110000	7	ACF65387_0	Acf65387 Photorhab
21	170.5	14.3	630	2	AAT67601	Aat67601 H. pylori
22	170.5	14.3	645	2	AAT68221	Aat68221 H. pylori
23	165	13.9	7521	5	AAS71378	Aas71378 DNA encod
24	165	13.9	7521	5	AAS94251	Aas94251 DNA encod
25	164	13.8	615	7	ACF73766	Acf73766 Staphyloc
26	154.5	13.0	3520	4	AAH54185	Aah54185 S. epider
27	135.5	11.4	110000	7	ACF67367_28	Continuation (29 o
c 28	135.5	11.4	110000	7	ACF65386_0	Acf65386 Photorhab
29	132.5	11.1	615	7	ACF69801	Acf69801 Photorhab
30	132.5	11.1	678	5	AAH65130	Aah65130 C glutami
31	132.5	11.1	678	7	ACA01893	Aca01893 C. glutam
32	132.5	11.1	349980	5	AAH64966	Aah64966 C glutami
33	121.5	10.2	708	8	ADA32836	Ada32836 DNA encod
34	121	10.2	405	2	AAV75577	Aav75577 Staphyloc
35	119.5	10.0	522	6	ABN91726	Abn91726 Staphyloc
36	119.5	10.0	669	8	ADA31058	Ada31058 DNA encod
37	119	10.0	1863	7	ACA53564	Aca53564 Prokaryot
c 38	116	9.7	110000	7	ACF65385_1	Continuation (2 of
c 39	116	9.7	110000	7	ACF67367_30	Continuation (31 o
c 40	114.5	9.6	1668	3	AAA61501	Aaa61501 A. vitis
41	114.5	9.6	1668	3	AAA61502	Aaa61502 A. vitis
42	112	9.4	609	7	ACF70009	Acf70009 Photorhab
43	108.5	9.1	636	7	ACF68771	Acf68771 Photorhab
44	108.5	9.1	717	8	ADA31709	Ada31709 DNA encod
45	108.5	9.1	110000	7	ACF67367_15	Continuation (16 o

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 7, 2004, 16:03:58 ; Search time 81 Seconds  
(without alignments)  
1616.894 Million cell updates/sec

Title: US-10-716-480A-2  
Perfect score: 1191  
Sequence: 1 MVIMEIFITGLLLGASLLLS.....INVVVAVVMTALAIKMLMG 236

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp

Q=/cgn2\_1/USPTO\_spool\_p/US10716480/runat\_06102004\_102310\_20256/app\_query.fasta\_1  
.391

-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10716480\_CGN\_1\_1\_105\_runat\_06102004\_102310\_20256 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	341	28.6	642	4	US-09-489-039A-4674	Sequence 4674, Ap
	2	332	27.9	699	4	US-09-252-991A-7908	Sequence 7908, Ap
c	3	332	27.9	834	4	US-09-252-991A-7643	Sequence 7643, Ap
	4	332	27.9	894	4	US-09-252-991A-7841	Sequence 7841, Ap
	5	327.5	27.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	6	327.5	27.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	7	318	26.7	597	4	US-09-894-844-11	Sequence 11, Appl
c	8	318	26.7	15239	1	US-08-390-878-17	Sequence 17, Appl
	9	310	26.0	645	4	US-09-543-681A-800	Sequence 800, App
	10	298	25.0	636	4	US-09-328-352-1611	Sequence 1611, Ap
c	11	157	13.2	1176	4	US-09-252-991A-7571	Sequence 7571, Ap
	12	138.5	11.6	639	4	US-09-543-681A-595	Sequence 595, App
	13	135.5	11.4	783	4	US-09-252-991A-9421	Sequence 9421, Ap
	14	135.5	11.4	1596	4	US-09-252-991A-9478	Sequence 9478, Ap
c	15	135.5	11.4	1650	4	US-09-252-991A-9352	Sequence 9352, Ap
	16	121.5	10.2	708	4	US-09-328-352-4123	Sequence 4123, Ap
	17	121	10.2	405	4	US-08-956-171E-1266	Sequence 1266, Ap
	18	119.5	10.0	522	4	US-09-134-001C-1189	Sequence 1189, Ap
	19	119.5	10.0	669	4	US-09-328-352-2345	Sequence 2345, Ap
	20	118.5	9.9	753	4	US-09-543-681A-315	Sequence 315, App
c	21	118.5	9.9	1713	4	US-09-252-991A-9760	Sequence 9760, Ap
	22	118.5	9.9	2805	4	US-09-252-991A-10208	Sequence 10208, A
	23	117.5	9.9	621	4	US-09-252-991A-2175	Sequence 2175, Ap
c	24	117.5	9.9	813	4	US-09-252-991A-1792	Sequence 1792, Ap
	25	115.5	9.7	660	4	US-09-489-039A-560	Sequence 560, App
	26	113.5	9.5	687	4	US-09-543-681A-682	Sequence 682, App
	27	111	9.3	669	4	US-09-252-991A-10073	Sequence 10073, A
	28	108.5	9.1	717	4	US-09-328-352-2996	Sequence 2996, Ap
	29	108	9.1	654	4	US-09-489-039A-905	Sequence 905, App
	30	104.5	8.8	636	4	US-09-543-681A-2536	Sequence 2536, Ap
	31	104	8.7	750	4	US-09-252-991A-9146	Sequence 9146, Ap
c	32	103	8.6	1650	4	US-09-489-039A-1643	Sequence 1643, Ap
	33	103	8.6	3324	4	US-09-489-039A-1719	Sequence 1719, Ap
	34	101.5	8.5	663	4	US-09-328-352-777	Sequence 777, App
	35	99	8.3	627	4	US-09-328-352-985	Sequence 985, App
	36	97.5	8.2	636	4	US-09-489-039A-2437	Sequence 2437, Ap
c	37	96	8.1	3402	4	US-09-252-991A-1374	Sequence 1374, Ap
	38	96	8.1	3687	4	US-09-252-991A-1193	Sequence 1193, Ap
	39	96	8.1	4266	4	US-09-252-991A-1234	Sequence 1234, Ap
	40	94.5	7.9	723	4	US-09-328-352-3411	Sequence 3411, Ap
	41	94	7.9	1659	4	US-09-489-039A-248	Sequence 248, App
	42	93	7.8	687	4	US-09-252-991A-7266	Sequence 7266, Ap
c	43	93	7.8	1521	4	US-09-252-991A-7437	Sequence 7437, Ap
	44	93	7.8	2397	4	US-09-252-991A-7203	Sequence 7203, Ap
c	45	91.5	7.7	7754	4	US-09-634-238-3	Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-09-489-039A-4674

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

```
Run on:      October 7, 2004, 16:11:23 ; Search time 407 Seconds
              (without alignments)
              2939.601 Million cell updates/sec
```

```
Title:          US-10-716-480A-2
Perfect score: 1191
Sequence:      1 MVIMEIFITGLLLGASLLLS.....INVVVAVVMTALAIKLMLMG 236
```

```
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
```

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Command line parameters:

```
-MODEL=frame+_p2n.model -DEV=xlp
```

Q=/cgn2\_1/USPTO\_spool\_p/US10716480/runat\_06102004\_102310\_20286/app\_query.fasta\_1  
.391

```
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10716480 @CGN_1_1_783 @runat_06102004_102310_20286
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

```
Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
```

11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*  
 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	% Query					Description
	No.	Score	Match Length	DB	ID	
	1	1191	100.0	708	9	US-09-738-626-3455 <i>NAK US</i> Sequence 3455, Ap
	2	1191	100.0	711	15	US-10-166-142-7 <i>NAK</i> Sequence 7, Appli
	3	1191	100.0	711	17	US-10-716-473-7 <i>NAK</i> Sequence 7, Appli
	4	1191	100.0	822	10	US-09-746-660A-51 <i>POMP US</i> Sequence 51, Appl
	5	1191	100.0	2374	15	US-10-196-232-24 <i>YOMA</i> Sequence 24, Appl
c	6	1191	100.0	3309400	9	US-09-738-626-1 <i>NAK</i> Sequence 1, Appli
	7	1169	98.2	712	15	US-10-166-142-9 <i>NAK</i> Sequence 9, Appli
	8	1169	98.2	712	17	US-10-716-473-9 <i>NAK</i> Sequence 9, Appli
	9	327.5	27.5	86114	15	US-10-080-170-648 Sequence 648, App
	10	327.5	27.5	86114	17	US-10-080-170-648 Sequence 648, App
	11	318	26.7	597	9	US-09-894-844-11 Sequence 11, Appl
	12	318	26.7	597	13	US-10-647-089-11 Sequence 11, Appl
	13	318	26.7	597	16	US-10-388-902-11 Sequence 11, Appl
	14	301	25.3	609	15	US-10-156-761-1109 Sequence 1109, Ap
c	15	301	25.3	9025608	15	US-10-156-761-1 Sequence 1, Appli
	16	170.5	14.3	630	13	US-10-335-977-1122 Sequence 1122, Ap
	17	170.5	14.3	633	13	US-10-335-977-1121 Sequence 1121, Ap
	18	170.5	14.3	645	13	US-10-335-977-1123 Sequence 1123, Ap
	19	133	11.2	672	15	US-10-156-761-246 Sequence 246, App
	20	132.5	11.1	678	9	US-09-738-626-165 Sequence 165, App
	21	132.5	11.1	3309400	9	US-09-738-626-1 Sequence 1, Appli
	22	128.5	10.8	32329	13	US-10-374-903A-1 Sequence 1, Appli
	23	121	10.2	405	8	US-08-781-986A-1266 Sequence 1266, Ap
	24	121	10.2	405	13	US-10-329-624-1266 Sequence 1266, Ap
	25	119	10.0	1863	13	US-10-282-122A-41434 Sequence 41434, A
	26	109.5	9.2	630	15	US-10-156-761-7444 Sequence 7444, Ap
	27	103.5	8.7	543	13	US-10-282-122A-8689 Sequence 8689, Ap
	28	99.5	8.4	888	13	US-10-282-122A-41598 Sequence 41598, A
	29	98.5	8.3	651	15	US-10-156-761-3368 Sequence 3368, Ap
	30	97.5	8.2	750	13	US-10-620-487-1 Sequence 1, Appli
	31	96.5	8.1	1018	13	US-10-425-114-28659 Sequence 28659, A
	32	94	7.9	1630	17	US-10-437-963-62711 Sequence 62711, A
	33	93.5	7.9	1359	9	US-09-815-242-6086 Sequence 6086, Ap
	34	93.5	7.9	1359	13	US-10-282-122A-20388 Sequence 20388, A
	35	93	7.8	633	13	US-10-282-122A-11642 Sequence 11642, A
	36	93	7.8	894	15	US-10-156-761-1180 Sequence 1180, Ap
	37	93	7.8	1377	9	US-09-738-626-3498 Sequence 3498, Ap
	38	93	7.8	9025608	15	US-10-156-761-1 Sequence 1, Appli

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

```
Run on:          October 7, 2004, 16:02:13 ; Search time 2437 Seconds  
                (without alignments)  
                2891.865 Million cell updates/sec
```

Title: US-10-716-480A-2  
Perfect score: 1191  
Sequence: 1 MVIMEIFITGLLLGASLLLS.....INVVVAVVMTALAIKLMLMG 236

```
Scoring table:  BLOSUM62
                  Xgapop 10.0 , Xgapext 0.5
                  Ygapop 10.0 , Ygapext 0.5
                  Fgapop  6.0 , Fgapext 7.0
                  Delop   6.0 , Delext  7.0
```

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
```

Q=/cgn2\_1/USPTO\_spool\_p/US10716480/runat\_06102004\_102310\_20237/app\_query.fasta\_1  
.391

```
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716480_CGN_1_1_5180_runat_06102004_102310_20237 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

```
Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
```

11: gb\_htc:\*  
 12: gb\_est3:\*  
 13: gb\_est4:\*  
 14: gb\_est5:\*  
 15: em\_estfun:\*  
 16: em\_estom:\*  
 17: em\_gss\_hum:\*  
 18: em\_gss\_inv:\*  
 19: em\_gss\_pln:\*  
 20: em\_gss\_vrt:\*  
 21: em\_gss\_fun:\*  
 22: em\_gss\_mam:\*  
 23: em\_gss\_mus:\*  
 24: em\_gss\_pro:\*  
 25: em\_gss\_rod:\*  
 26: em\_gss\_phg:\*  
 27: em\_gss\_vrl:\*  
 28: gb\_gss1:\*  
 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result			%		Query		DB	ID	Description
	No.	Score	Match	Length					
c	1	317.5	26.7	830	29	CNS01MQH			AL151258 Anopheles
c	2	281	23.6	667	28	AQ991044			AQ991044 Rfc01899
	3	207.5	17.4	1034	28	BZ554294			BZ554294 pacs1-60_
c	4	205.5	17.3	833	28	BZ556963			BZ556963 pacs1-60_
c	5	199	16.7	804	29	CNS01HIJ			AL144476 Anopheles
c	6	160	13.4	897	28	BZ560550			BZ560550 pacs2-164
c	7	149	12.5	1620	28	BZ568946			BZ568946 pacs2-164
c	8	143	12.0	751	28	BH391509			BH391509 AG-ND-159
	9	116.5	9.8	909	28	BZ565158			BZ565158 pacs2-164
	10	98	8.2	829	28	CC113281			CC113281 NDL.24H1.
	11	94.5	7.9	1525	29	AY415555			AY415555 Pan trogl
c	12	94	7.9	840	29	CC719070			CC719070 OGWCI79TV
	13	93.5	7.9	210	14	CB020010			CB020010 pw98f01.y
	14	93.5	7.9	570	14	CF275208			CF275208 NcEST3d13
	15	93.5	7.9	741	13	BX612630			BX612630 BX612630
	16	91	7.6	632	12	BI862608			BI862608 603389677
c	17	90	7.6	798	29	CNS01QYA			AL156722 Anopheles
	18	89.5	7.5	679	13	BW283758			BW283758 BW283758
	19	89	7.5	590	14	CD482346			CD482346 atr01-4ms
	20	89	7.5	709	14	CA300534			CA300534 SCSFLV104
	21	88.5	7.4	438	28	AQ856877			AQ856877 nbeb0004K
c	22	88.5	7.4	945	13	BU554436			BU554436 AGENCOURT
c	23	88.5	7.4	1233	28	BZ579504			BZ579504 msh2_6385
c	24	87.5	7.3	628	14	CA279614			CA279614 SCCCFL800
	25	87.5	7.3	656	13	BU002594			BU002594 QGG31M20.
c	26	87.5	7.3	682	29	CC968186			CC968186 BOICV59TR
	27	87.5	7.3	684	13	BQ856142			BQ856142 QGB28L08.
	28	87.5	7.3	710	13	BQ866753			BQ866753 QGC8N13.y